Knowledge Acquisition from Clinical Case Reports: Quality and Utility for Case-based Biomedical Article Retrieval

Matthew S. Simpson, MSc; Charles Sneiderman, MD, PhD; Dina Demner-Fushman, MD, PhD; and George R. Thoma, PhD

Lister Hill National Center for Biomedical Communications U.S. National Library of Medicine, Bethesda, MD 20894

ABSTRACT

Case-based article retrieval refers to the task of automatically finding clinical case reports that are similar to a given topic case. A crucial step in this process is the extraction and representation of the most salient information from a topic case in order to create a well-formed clinical query. In this study, we report on the quality of automatically extracted terms and their utility in facilitating case-based article retrieval.

OBJECTIVES & METHODS

Case-based retrieval is an essential part of our research into providing multimedia evidence to support clinical decision making.[1] In this study, we evaluate our current approach to formally representing case descriptions and automatically generating queries based on this structured representation.

We focus our experiment on case descriptions from the *Journal of Family Practice* "Photo Rounds" feature. A "Photo Rounds" article presents a detailed description of a clinical case and describes a differential diagnosis with supporting evidence from references. Our task is to retrieve similar cases by forming a query from terms automatically extracted from the case description.

In constructing our test document collection, we assumed an article's cited references to be the set of *relevant documents* we aimed to retrieve. Additionally, to add noise to the collection, we included articles from various sources, each with *no relevance judgement*.

We formally represented case descriptions following the well-formed clinical question framework using extracted UMLS® terms related to problems, interventions, patient characteristics, and image modality as described in our previous work.[1,2] We formed each query as the disjunction of all non-negated terms.

To evaluate the *quality* of our method, we enlisted a family physician trained in informatics to manually annotate terms if they were "useful" for constructing a clinical query, "correct" if they were identified with the correct UMLS semantic type and negation status, or "both." The annotator also added additional useful terms that were not extracted. To evaluate their *utility*, we performed batch queries for each case description using (1) all extracted terms, (2) only useful terms, (3) only correct terms, and (4) terms that were both useful and correct. We report precision and recall for extraction quality and bpref[3] for utility in retrieval.

Used Terms	Extraction		Retrieval	
	Precision	Recall	bpref	δ (%)
All extracted	–	–	0.7379	–
Only useful	0.8349	0.7622	0.7921	7.343
Only correct	0.8372	0.7602	0.7921	7.343
Useful & correct	0.7756	0.7476	0.7921	7.343

Table 1: Quality and utility of extracted terms.

RESULTS

Table 1 summarizes the quality and utility of the extracted terms. For identifying useful terms, our extractor achieved a precision of 0.83 at 0.76 recall. In retrieving relevant cases, we saw a moderate 7.3% improvement (δ) in bref ($p < 0.05$) when eliminating terms that were not useful or correct. This indicates our approach to case-based retrieval can benefit from an improved term extraction method.

REFERENCES

1. Simpson M, Rahman MM, Demner-Fushman D, Antani S, Thoma GR. Text- and content-based approaches to image retrieval for the imageclef 2009 medical retrieval track; 2009.
2. Demner-Fushman D, Antani S, Simpson M, Thoma GR. Annotation and retrieval of clinically relevant images. Int J Med Inform 2009;78(12):e59–67.
3. Buckley C, Voorhees EM. Retrieval evaluation with incomplete information. In: Proceedings of the 27th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval. Sheffield, United Kingdom: ACM; 2004, p. 25–32.